Churn Management Optimization with Controllable Marketing Variables and Associated Management Costs

ABSTRACT

In this paper, we propose a churn management model based on a partial least square (PLS) optimization method that explicitly considers the management costs of controllable marketing variables for a successful churn management program. A PLS prediction model is first calibrated to estimate the churn probabilities of customers. Then this PLS prediction model is transformed into a control model after relative management costs of controllable marketing variables are estimated through a triangulation method. Finally, a PLS optimization model with marketing objectives and constraints are specified and solved via a sequential quadratic programming method. In our experiments, we observe that while the training and test data sets are dramatically different in terms of churner distributions (50% vs. 1.8%), four controllable variables in three marketing strategies significantly changed through optimization process while other variables only marginally changed. We also observe that the most significant variable in a PLS prediction model does not necessarily change most significantly in our PLS optimization model due to the highest management cost associated, implying differences between a prediction and an optimization model. Finally, two marketing models designed for targeting the subsets of custo mers based on churn probability or management costs are presented and discussed.

Keywords: Churn management, PLS prediction model, PLS optimization model, Management cost, Triangulation, Sequential quadratic programming;

1. INTRODUCTION

The propensity of customers to terminate their relationships with service providers has forced many companies in competitive markets to shift their strategic focus from customer acquisition to customer retention (Chen and Hitt, 2002; Venkatesan and Kumar, 2004). This is mainly because companies can increase the average net present value of a customer by up to 95% by boosting the customer retention rates by 5%. In particular, with exceptionally high annual churn rates (20–40%), the mobile telecommunications service providers eager to launch successful churn management programs to maximize their revenues (Kim et al., 2004; Eshghi, et al., 2007; Glady et al., 2009). For this purpose, many data mining and statistical models have

been presented to accurately identify prospects or possible churners in the automotive, insurance, and telecommunication industry. Such models include PLS regressions (Lee et al., 2011), decision trees (Kim, 2006; Xiea et al., 2009), ANNs (Mozer et al., 2000; Buckinx and Poel, 2005), support vector machines (Coussement and Poel, 2008), genetic algorithms (Au et al., 2003; Kim et al., 2005), or dynamic programming models (Gönül and Shi, 1998). A recent discussion about advantages and disadvantages of various models for churn management can be found in (Neslin et al., 2006; Hadden et al., 2007).

While such prediction models are very important for successful churn management, most prediction models are limited in the sense that they do not consider implementations costs associated with churn management programs (Bult and Wansbeek, 1995; Kumar and Shah, (2004). Note that, according to these predictive models, marketing managers may identify likely churners based on the estimated churn probability, and chooses top x% of customers as target customers to offer retention marketing promotions (Lee et al., 2011). However, most retention marketing promotions bear different cost structures and hence should be administered with care. For example, one of the most common retention programs among telecommunication service providers is to provide customers who renew their contract periods a financial incentive that allows them to purchase a new mobile device at a deeply discounted price. Another popular retention program is to simply provide a better customer call center service in regards to billing and call quality peacefully through educated and experienced receptionists to resolve many questions and complaints and enhance customer satisfaction and loyalty (Fornell and Wernerfelt, 1987; Mittal and Kamakura, 2001; Reinartz et al., 2005; Gustafsson et al., 2005). Note that while two retention programs may or may not be equally effective, providing a new mobile device at a deeply discounted price may cost more than providing a better call center service.

In this paper, we propose a churn management model based on partial least square (PLS) optimization that explicitly considers management costs of controllable marketing variables. The PLS method in this paper will be used not only as a prediction model to predict churners but also as a control model combined with optimization method to maximize the effects of churn management strategies at minimum cost. Ideally, limited resources for retention promotions should be allocated to most likely churners who generate most revenues while minimizing the management costs of such retention promotions. The detailed objectives of this research are: (1) to categorize and validate controllable and uncontrollable marketing variables; (2) to determine

the management costs of each controllable marketing variable by applying a triangulation analogy method; and (3) to develop and solve a PLS optimization model that minimizes the total cost of implementing three retention marketing strategies while satisfying the objective of retention marketing strategies.

The remainder of this paper is organized as follows. Section 2 provides a brief review of PLS model for prediction and control purposes and a triangulation method for management cost estimation. In Section 3, the overall research framework is introduced, and data sets are explained. In the following Section 4, a PLS-based optimization model is presented in a mathematical form after controllable marketing variables are identified and their management costs are assigned. Section 5 first presents experimental results from a PLS optimization model designed for entire customers. Then two marketing models designed for targeting the subsets of customers based on churn probability or management costs are presented and discussed. Finally, Section 6 provides the conclusion of the paper and suggests several direction of further research.

2. PARTIAL LEAST SQUARE AND TRIANGULATION ANALOGY

2.1 Partial least square (PLS) method

The PLS method is a multivariate projection approach that can consider both multiple responses and multiple predictors variables. In particular, it has been known to be robust with data sets that contain measurement errors and collinearity (Geldadi and Kowalski, 1986; Malthouse et al., 1997). Naturally, one of the most popular applications of PLS models is to transform original large-scale data into lower dimensional data to deal with highly correlated data between independent and dependent variables (Lakshminaraynan et al., 1997). In this process, several PLS factors are extracted to explain most of the variation in both independent and dependent variables (Chong et al., 2007). When a nonlinear relationship is implicitly assumed among dependent and independent variables, nonlinear PLS models can be estimated to construct nonlinear functional relationships using either neural networks or Gaussian kernel (Qin and McAvoy, 1992; Malthouse et al., 1997; Rosipal and Trejo, 2002; Shawe-Taylor and Cristianini, 2004).

The PLS method can be valuable as an alternative to well known data mining models to predict response variables or as a path model to understand structural relationships among records. In a recent study (Laitinen, 2008), PLS is utilized in building a predictive system with

some success to assess failure probability in small- and medium-sized Finnish firms using financial and non-financial variables and reorganization plan information. In several other studies (Qin and McAvoy, 1992; Wiener et al., 2010; Lee et al., 2011; Kim et al., 2012), PLS prediction models not only showed superior or comparable performance against other data mining algorithms, but also showed the usefulness of identifying key input variables for biology and marketing applications. However, we have not seen papers that apply a PLS model both as a prediction model and as an optimization model for churn marketing management. Note that a PLS optimization model can be combined with a PLS prediction model as well as any one of data mining prediction algorithms. Since several comparative studies (Lee et al., 2011; Kim et al., 2012; and references therein) investigated the performance of PLS prediction models against different prediction models, we limit our interests to the specification and application of a PLS optimization model assuming that a (PLS) prediction model is readily available.

In this study, a linear PLS method (Lee et al., 2011) is first used as a prediction model to predict churners based on demographic, psychographic, and historical service usage information. Note that in prediction tasks, the number of latent variables is an important factor that affects the predictive accuracy. Typically, the number of latent variables is chosen by a cross validation considering the proportion of variations explained by each latent variable. At the same time, the variable importance in projection (VIP) has been proposed and used as a measure of importance of each variable contributing the response of interest. The VIP of *j*-th variable is calculated as the sum of weights on the latent variable from *j*-th variable divided by the total weights. Fundamentally, our final PLS prediction model includes only a set of predictors with high VIP scores to improve the comprehensibility and reduce computational complexity of a prediction model.

The PLS method is also used as an optimization model for churn management in our study. For churn management optimization, the PLS optimization model is first specified after marketing managers identify controllable and uncontrollable marketing variables among the chosen input variables with high VIP scores from the PLS prediction model. Then managers subjectively assign to controllable marketing variables management costs that are required to change one unit value of the chosen controllable marketing variable. Once the overall marketing objective (e.g., reducing churn probabilities of all the customers on average by 20% while minimizing the management costs of marketing variables) and constraints are specified, the PLS

optimization model can be solved with an iterative optimization algorithm procedure such as sequential quadratic programming technique.

2.2 Triangulation method for estimating management costs

The main objective of this study is to reduce churn probabilities of all the customers on average by 20% while minimizing the management costs of controllable marketing variables. Therefore, to find realistic and meaningful solutions from the PLS optimization model for churn management, it is necessary to assign the management costs to identified controllable variables. While it is ideal to use the actual implementation costs of controllable variables in the PLS optimization model, it is extremely difficult to estimate the actual costs of marketing variables due to the lacks of historical data and/or the inseparability of associated indirect and direct costs. Therefore, it is necessary to use a new method to estimate the management costs of controllable marketing variables and a triangulation method that has been popularly adopted in agile software engineering and development community (Cohn, 2006) can be regarded as an attractive and feasible alternative. Note that one of the most important and difficult tasks of the project managers in software development projects is to accurately estimate the size and duration of projects, which in turn determines the scheduling and budgets.

The triangulation method is based on the fact that humans are better at estimating relative size than estimating absolute size (Lederer, 1998; Vicinanza et al., 1991). According to this observation, we can easily tell which animal is the tallest in the ascending order from a set of animals (e.g., a mouse, a cat, a tiger, and an elephant) while we do not know exactly how tall each animal is. In addition, humans have been known to be good at estimating things that fall within a single order of magnitude (Miranda, 2001), implying that it is fairly easy to estimate the relative distance from home to the nearest grocery store compared with the distance to the nearest library (e.g., the library may be twice as far as the nearest grocery) while it is difficult to estimate the relative distance to the neighboring country's capital.

Based on the afore-mentioned discussions, we like to estimate and assign to controllable marketing variables management costs that reflect "relative difficulty" of controlling a specific variable compared with other variables. For this purpose, one of two popular nonlinear estimation scales, Fibonacci sequence (1, 2, 3, 5, and 8), is adopted as the range of possible estimates of difficulty. To apply this method, a marketing manager may select one of the "easiest" variables to

control among all the available marketing variables and assign the lowest value of "1" to this variable. Then, for each remaining controllable variables, its management cost is determined by comparing its managerial difficulty against that of the easiest variable (i.e., one with the cost estimate of "1") or an assortment of those that have already been estimated. Then, the closest value from the Fibonacci sequence is chosen as the final estimate. For example, if a chosen marketing variable is twice more difficult to manage than others, its management cost is estimated to be twice as others.

Note that this method based on Fibonacci sequence naturally reflects the greater risks associated with estimates for much more difficult marketing variables because of its nonlinear sequence property. This method is very attractive because the management costs can be prepared by any marketing manager with unique preferences and experiences who will be responsible for the final results of marketing campaign (Lederer and Prasad, 1992). For even better results, this method can be also implemented with a set of marketing managers through a planning poker game where the cost estimates of all marketing managers who estimate management costs by using the same deck of cards of Fibonacci sequence are averaged to obtain better results (Johnson et al., 2000). This way, the unique preferences and experiences of each marketing manager are accommodated to accurately estimate the management costs of controllable variables with high uncertainty (Hagafors, 1983) through discussion among managers with high and low estimates. Further, this method naturally reflects the greater risks associated with estimates for difficult variables because of nonlinear property of Fibonacci sequence.

3. RESEARCH FRAMEWORK AND DATA SET

3.1 Research framework

Our research framework is illustrated in Figure 1. As the first step, it is necessary to preprocess raw data into a readily available format for further analysis. In this study, two different techniques—eliminating records with missing values and variable selection—are used separately or together for preprocessing raw data. Once preprocessed data sets are obtained, a PLS prediction model is developed to estimate the churn probability for each customer. Note that our PLS prediction model is built with a set of chosen input variables with high VIP values and several latent variables (i.e., score variables) only to reduce computational complexity and improve the managerial understanding of outcomes.

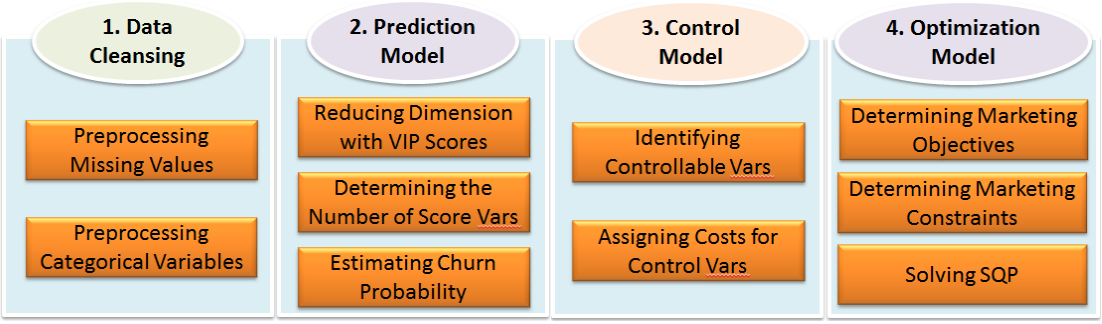

Figure 1. Research framework

At the third step, a PLS control model is conceptually developed. Marketing managers can use this model to accomplish enterprise marketing objectives by controlling key marketing instruments. To develop a PLS control model, we first categorize variables in a PLS prediction model into controllable and uncontrollable variables and assign management costs for each controllable variable. Note that input variables that are useful for building a highly accurate prediction model may not be ideal for a PLS control model when they are not controllable or are associated with high management costs. Once the PLS control model is specified, it is converted into an optimization model by specifying the marketing objective in a mathematical expression. If necessary, we limit the number of changes and ranges of controllable marketing instruments in the objective function and additional constraints. Finally, the designed optimization problem is solved using one of iterative optimization algorithms (e.g., successive quadratic programming (SQP) (Biegler, 1995)), which transforms the original problem into an easier sub-problem that can be easily solved and then used as a basis of iterative procedure.

3.2 Data set

The data sets used in this study are provided by the Teradata Center for CRM at Duke University, and the original data set for calibration has 171 predictor variables of 100,000 observations. The complete set of variables includes three types of variables: behavioral information such as minutes of use, revenue, handset equipment; company interaction information such as customer calls into the customer service center; and customer household demographics. For each customer, churn was calculated based on whether the customer left the company during the 31- to 60-day period after the customer was originally sampled. Although the actual percentage of customers who left the company in a given month is approximately 1.8%, churners in the original data set were oversampled to create roughly a 50–50 split between

churners and nonchurners. However, the test data set with 51,306 observations are expected to represent a realistic churning rate, 1.8%.

As a preprocessing step for further analysis, the original data sets are preprocessed as follows. First, most categorical variables are excluded because of high missing rate or being encoded into multiple binary variables which makes low predictive power. We include only 11 categorical variables which are either indicator variables or countable variables such as number of handsets and number of subscribers. This is because each categorical variable has very little predictive power in general (Rossi et al., 1996). Second, continuous variables with more than 20% of missing values are eliminated. We take 123 predictors including 11 categorical variables and 112 continuous variables in data preprocessing step. Finally, records with missing values in the data set with 123 predictors are removed from further analysis. After preprocessing steps, the training set contains 67,181 observations with 32,862 churners (48.92%), while the test set contains 34,986 observations with 619 churners (1.8%), respectively.

4. DESIGN OF PLS PREDICTION, CONTROL, AND OPTIMIZATION MODEL

4.1 Design of PLS prediction model

The main objective of constructing a PLS prediction model is to accurately classify churners against non-churners or estimate the churning probability of all the service users. The design process of a PLS prediction model is a fairly straight forward process and it is similar to the process of building a typical regression model. However, an initial PLS prediction model includes all the input variables and hence it may not be scalable, generalizable, and interpretable. To further enhance the scalability and interpretability of the initial PLS prediction model, the VIP score of each variable is often computed and only input variables that meet a certain threshold are included in the final PLS prediction model. After we explored a set of VIP cut-off criteria (0, 1.0, 1.2, and 1.5), we found that the PLS prediction model with a VIP cut-off criterion of 1.0 provided a reasonably good compromise of the most parsimonious PLS model (i.e., a model with the highest VIP cut-off criterion of 1.5) and the most comprehensive model (i.e., a model with all the input variables or a model with the lowest VIP cut-off criterion of 0). Therefore, our final PLS prediction model is based on 46 variables (out of 123 original variables) with high VIP scores greater than or equal to 1.0 as shown in Table 1 and it includes six latent variables that is a linear (or nonlinear if nonlinear PLS model is adopted) combination of 46 variables.

Table 1. Variables chosen in the final PLS prediction model (Lee et al., 2011)

Rank	Variable Description	VIP Score	Rank	Variable Description	VIP Score
1	Number of days (age) of current equipment	2.995	24	Mean monthly revenue (charge amount)	1.2278
2	Handset: refurbished or new	1.9089	25	Mean rounded minutes of use of customer care calls	1.1984
3	Age of first household member	1.7519	26	Number of models issued	1.1964
4	Average monthly minutes of use over the life of the customer	1.5625	27	Range of unrounded minutes of use of peak voice calls	1.182
5	Range of revenue of voice overage	1.5562	28	Mean unrounded minutes of use of customer care calls	1.1652
6	Range of revenue of overage	1.5556	29	Average monthly revenue over the previous three months	1.1572
7	Percentage change in monthly minutes of use vs. previous three month average	1.5406	30	Average monthly revenue over the life of the customer	1.1497
8	Range of overage minutes of use	1.5302	31	Mean number of customer care calls	1.1307
9	Average monthly number of calls over the life of the customer	1.5211	32	Total number of calls over the life of the customer	1.1162
10	Mean revenue of voice overage	1.4597	33	Mean number of monthly minutes of use	1.1101
11	Mean overage revenue	1.4586	34	Number of unique subscribers in household	1.1094
12	Account spending limit	1.4448	35	Billing adjusted total number of calls over the life of the customer	1.1091
13	Mean overage minutes of use	1.4039	36	Mean number of dropped voice calls	1.0683
14	Range of revenue (charge amount)	1.3976	37	Average monthly number of calls over the previous six months	1.055
15	Total number of months in service	1.3268	38	Total minutes of use over the life of the customer	1.0485
16	Mean total monthly recurring charge	1.3051	39	Range of rounded minutes of use of customer care calls	1.0447
17	Number of handsets issued	1.3	40	Range of number of received voice calls	1.0445
18	Range of number of minutes of use	1.2882	41	Billing adjusted total minutes of use over the life of the customer	1.0375
19	Range of number of attempted voice calls placed	1.2703	42	Average monthly minutes of use over the previous three months	1.0375
20	Range of number of completed voice calls	1.2666	43	Range of unrounded minutes of use of customer care calls	1.0351
21	Range of number of attempted calls	1.2655	44	Range of total monthly recurring charge	1.0299
22	Range of number of completed calls	1.2597	45	Range of unrounded minutes of use of completed voice calls	1.0221
23	Range of number of inbound and outbound peak voice calls	1.2536	46	Range of number of off-peak voice calls	1.0036

The final prediction model also includes the impact direction of each variable ('+' implying a positive impact and '-' implying a negative impact) on the churning decision of service users. For example, the "Number of days (age) of current equipment" variable is associated with the highest VIP score (2.995) and positively affects the churning decision (i.e., a

service user who has kept her current mobile equipment longer is more likely to churn). It is also noted that the variable, "Handset-refurbished or new", is considered the second most important variable (VIP score = 1.9089) although it negatively affects the churning decision of service users. Note that while the number of selected variables was determined using VIP scores, the number of latent variables was determined using the cumulative variation explained by the latent variables. We chose six latent variables that explain 90% of the variable in the data. According to Table 2, as we sequentially add an additional latent variable, the cumulative percentage of explained variance in the dependent variable is increasing at a decreasing rate and the contributions from each additional latent variable after the first six variables becomes marginal (less than 1%).

Table 2. Variance explained by each latent variable

LV #	Variance Explained	Cumulative Variance Explained
1	37.8%	37.8%
2	14.61%	52.42%
3	27.8%	80.21%
4	8.17%	88.38%
5	1.32%	89.70%
6	0.39%	90.09%

4.2 Identification of controllable variables

Once the PLS prediction model is identified, we convert the prediction model into an optimization model by identifying controllable variables along with associated management costs and other constraints, and solve it using a SQP algorithm. The first step then is to first formulate the PLS control model by dividing the variables in the PLS prediction model into controllable and uncontrollable variables, and assigning the associated costs to controllable variables. Note that not all predictive variables are controllable for optimization. To identify controllable variables out of 46 chosen marketing variables in Table 2, we assume that all the controllable variables can be adopted for one of three possible marketing strategies for churn management (Lee et al., 2011): device management strategy (DMS), revenue management strategy (RMS), and complaints management strategy (CMS). The main purpose of DMS is to manipulate controllable variables that are directly or indirectly related to mobile devices. A plausible example of DMS is to provide new mobile devices at deeply discounted prices for the customers whose service contracts will expire soon on the condition that they will renew their service

contracts for two more years. After carefully reviewing 46 input variables in the PLS prediction model, we identify five variables for DMS. These variables include “Number of days (age) of current equipment”, “Handset: refurbished or new”, “Total number of months in service”, “Number of handsets issued”, and “Number of models issued”. We conclude that all these variables except “Total number of months in service” can be controlled by the service provider directly by regulating the number of models or handsets issued to the customers or indirectly by providing strong price incentives to purchase new or refurbished handsets. We still consider “Total number of months in service” controllable if the service provider makes the customers satisfy with the current services.

The main purpose of OMS is to effectively control revenue related variables after carefully reviewing fee structures of various services and services usage patterns of customers. For example, one of possible OMS is to solicit users who frequently overuse their voice and data plan beyond their allowance for premium services with a marginal increase of service fees. This strategy will benefit heavy users because they enjoy upgraded services with a marginal fee increase and avoid unexpected and expensive overage charges. The service providers also benefit because they can turn uncertain cash flow from monthly overage charges into steady and predictable cash flow. From the PLS prediction model, we identify 10 variables related to OMS. These variables include “Range of revenue of voice overage”, “Range of revenue of overage”, “Mean revenue of voice overage”, “Mean overage revenue”, “Range of total monthly recurring charge”, “Mean total monthly recurring charge”, “Range of revenue (charge amount)”, “Mean monthly revenue (charge amount)”, “Average monthly revenue over the previous three months”, and “Average monthly revenue over the life of the customer”. Note that while the service providers cannot directly manipulate simple statistics such as “range”, “mean”, or “average” values of these variables, they can indirectly change these distributions of revenues through various OMS strategies, resulting in the changes of related statistics.

Finally, the main objective of CMS is to provide highly responsive services to the customers who have complained technical difficulties and billing discrepancies so that they are satisfied and decide to stay with the current service plan for the remaining contract period. Note that the successful implementation of CMS will significantly reduce the work loads of operators at the call centers, resulting in reduced operating costs (typically $2 per complain call). From the PLS prediction model, we identify three CMS related controllable variables such as “Account

spending limit [on or off]", "Billing adjusted total number of calls over the life of the customer", and "Billing adjusted total minutes of use over the life of the customer". Note that the service provider may not directly control the billing adjusted total number of calls or the billing adjusted total minutes of use because it is customers who decide to make claims. However, the service provider can definitely reduce billing related calls not only by improving the accuracy, reliability, and availability of its fee structures but also by maintaining informal sessions (e.g., FAQ on its Web site) about its unique fee structures with customers. In the end, a total of 18 controllable variables for DMS, OMS, and CMS are identified.

4.3 Assignment of management costs to controllable variables

Once the controllable variables associated with three marketing strategies are identified, the next step is to assign management costs to controllable variables. We followed the planning poker procedure explained in Section 2.2 using a Fibonacci sequence (1, 2, 3, 5, and 8) to assign the relative difficulty of controlling variables in three marketing strategies.

We first note that the implementing DMS based on device related controllable variables is relatively difficult without the mutual agreements between the service provider and customers. For example, the service provider cannot enforce customers to buy a new (or refurbished) device or to change the current device into a new model unless the customers want to do so, although it may influence the decision of customers by providing strong financial incentives to purchase new handsets. Further, the service provider can only control these variables after carefully considering its own and competitors' prices of devices. Therefore, we decide to assign high values (>= 5) as management costs to these controllable variables. In particular, we assign the highest management cost (i.e., the value of 8) to "Total number of months in service" variable because of its dependency on service quality and customer satisfaction in addition to device factors. In contrast, we assign the lowest value (i.e., the value of 1) to all CMS variables because the service provider can easily change them (e.g., for "Account spending limit" variable, the service provider can simply set on or off the spending limit of the chosen customer account). Finally, we assign from low to high values to the RMS variables. Specifically, we assign the value of 1 to all range variables of (overage) revenues because of their simple statistic natures, and the value of 3 to most mean or average of revenue variables. However, we assign the value of 5 to "Mean monthly

revenue (charge amount)" because of its dependency on service fee structures associated with customers' credit history and competitors' fee strategy.

We note that this process is necessarily based on subjective judgments of researchers or marketing managers who estimate the relative difficulty of manipulating controllable variables. While this can be regarded as a weakness of the proposed approach, it can be very useful because it allows any marketing managers in their unique organizational and business environments to reflect their own preferences and experiences, and to maximize the outcomes of their marketing programs for churn management. Further, our research framework is robust enough to accommodate various management cost structures of controllable variables and can still provide valid marketing insights.

4.4 Optimization model and solution procedure

Assuming that a marketing manager wants to reduce churn probabilities of all the customers on average by 20% by manipulating p controllable marketing variables while minimizing the management costs of controlling these variables, we formulate our PLS optimization model as follows:

$$\min_x \sum_{j=1}^{p} c_j * (x_j - x_j^{opti.})^2 \qquad (1)$$

$$\text{s.t. } \sum_{j=1}^{p} b_j * x_j^{opti.} \leq (0.8) * (\sum_{i=1}^{p} b_j * x_j) \qquad (2)$$

$$x_i \geq 0 \text{ for all } i \ , \ x_{aslflg}, x_{refurb} \leq 1 \qquad (3)$$

where x_j and $x_j^{opti.}$ represents the current and the desired value of a controllable variable j. Therefore, the objective function in Equation (1) represents the penalty on the deviations of x_j from $x_j^{opti.}$ weighted by c_j, the management cost of a controllable variable j. The constraint equations (2) and (3) simply specify the 20% reduction of churn probability associated with PLS regression coefficient b_j and feasible ranges of variables with the upper bounds of two indicator variables, "Accounting spending limit" and "Handset: refurbished or new", respectively. This non-linear optimization is then solved using the iterative optimization algorithms such as sequential quadratic programming (SQP) (Gill et al., 1984) which transforms the original problem into an easier sub-problem that can be easily solved and then used as a basis of iterative procedure. The SQP is well known for its efficiency, accuracy, and percentage of successful

solutions over a large number of problems (Schittowski, 1985).

5. EXPERIMENTAL RESULTS

5.1 Optimization model designed for entire customers

The outputs of the optimization model in terms of the mean values of controllable variables for all the customers before and after optimization from both training and test data sets are presented in Table 3. Note that, in Table 3, $Mean^{bo}$ represents the mean values of controllable marketing variables before optimization while $Mean^{ao}$ represents the mean values of optimized controllable marketing variables to reduce churn probabilities of all the customers on average by 20%.

Table 3. Mean values of controllable variables before and after optimization

Marketing Strategies	Variables	VIP Score	Churn Impact	Cost	Train data			Test Data		
					$Mean^{bo}$ 1)	$Mean^{ao}$ 2)	Change (%)	$Mean^{bo}$	$Mean^{ao}$	Change (%)
DMS	Number of days (age) of current equipment	2.995	+	5	415.51	415.31	-0.048	388.33	388.16	-0.044
	Handset: refurbished or new	1.9089	-	5	0.85538	0.9374	**9.589** 3)	0.86477	0.92674	**7.166**
	Total number of months in service	1.3268	-	8	19.915	21.482	**7.868**	19.805	21.096	**6.519**
	Number of handsets issued	1.3	+	5	1.794	0.87908	**-50.999**	1.8426	0.92185	**-49.970**
	Number of models issued	1.1964	-	5	1.5514	3.7318	**140.54**	1.5895	3.3868	**113.073**
RMS	Range of revenue of voice overage	1.5562	+	1	28.067	27.873	-0.691	26.414	26.263	-0.572
	Range of revenue of data overage	1.5556	+	1	28.511	28.303	-0.730	26.887	26.724	-0.606
	Mean revenue of voice overage	1.4597	+	3	12.134	12.067	-0.552	11.365	11.313	-0.458
	Mean overage revenue	1.4586	+	3	12.365	12.286	-0.639	11.616	11.554	-0.534
	Range of revenue (charge amount)	1.3976	+	1	39.983	39.392	-1.478	38.565	38.077	-1.265
	Mean total monthly recurring charge	1.3051	-	2	44.825	47.035	**4.930**	46.636	48.457	**3.905**
	Mean monthly revenue (charge amount)	1.2278	+	5	56.201	56.14	-0.109	56.892	56.842	-0.088
	Average monthly revenue over the previous three months	1.1572	-	3	56.528	56.988	0.814	57.448	57.827	0.660
	Average monthly revenue over the life of the customer	1.1497	+	3	54.823	53.921	-1.645	55.483	54.739	-1.341
	Range of total monthly recurring charge	1.0299	-	1	6.6881	10.863	**62.423**	7.3618	10.803	**46.744**
CMS	Account spending limit	1.4448	-	1	0.09685	0.94427	**874.90**	0.11522	0.9819	**752.196**
	Billing adjusted total number of calls over the life of the customer	1.1091	-	1	2792.4	2792.4	0.000	2856.6	2856.6	0.000
	Billing adjusted total minutes of use over the life of the customer	1.0375	-	1	7303.4	7303.4	0.000	7557.4	7557.4	0.000

1) $Mean^{bo}$: mean values of controllable variables before optimization
2) $Mean^{ao}$: mean values of controllable variables after optimization
3) Changes (%) in bolds were statistically significant at $\alpha = 0.01$

We immediately notice that the changes in the mean values of all the variables during the optimization were consistent with our expectation: mean values of marketing variables that positively (negatively) impact on churn decision decrease (increase) to reduce churn probabilities. One interesting finding is that while the training and test data sets are dramatically different in terms of churner distributions (50% vs. 1.8%), the same seven variables in bold fonts were statistically significantly changed ($p < 0.001$) while other variables were not changed at all or only marginally changed. In particular, four variables such as two DMS variables ("Number of handsets issued" and "Number of models issued"), "Range of total monthly recurring charge" from RMS, and "Account spending limit" from CMS most significantly changed (i.e., higher than 40%) during the optimization. This indicates that these four controllable variables from three marketing strategies should be considered simultaneously for optimization to maximize the synergistic effects of multiple marketing strategies. In contrast, two CMS controllable variables ("Billing adjusted total {number, minutes} of calls over the life of the customer") do not change at all during the optimization process due to their limited impacts on the completion of optimization objective specified in Equation (1).

Another interesting finding is that the most important two variables for prediction in terms of VIP scores do not necessarily change most significantly during the optimization process. For example, the change of "Number of days (age) of current equipment" variable with the highest VIP score was not statistically significant. While the change of "Handset: refurbished or new" variable with the second highest VIP score (7.166% on test set) was statistically significant, but the magnitude of its change was minimal when compared with changes in afore-mentioned four variables. These finding imply that each marketing variable plays different role in prediction and optimization models, and this warrants further investigations.

5.2 Optimization model designed for subsets of customers

5.2.1 Customer selection based churn probability

The discussion in Section 5.1 focused on the changes in the mean values of controllable variables over all customers to reduce their churn probabilities by 20%. However, not all customers have the same probability of switching the current service provider. In addition, the management cost to reduce the churn probability of each customer can be significantly different. The Figure 2 shows that the finalized values of controllable variables to reduce churn

probabilities of each customer by 20% along with associated management costs and churn probabilities before and after optimization. For example, the customer (ID: 2000004) is very likely to churn (churn probability before optimization is 75.26%) and requires a high management cost (142.86) to reduce her churn probability by 20% (from 75.26% to 60.21%) by changing the values of controllable variables as shown in the corresponding row in the left side of the figure. In contrast, the customer (ID: 2000009) is associated with a very low churn probability and a very low management cost (0.25365). In fact, customers with high estimated churn probabilities are associated with high churn management costs mainly because it is expected to adjust the values of controllable variables "more significantly" for these customers during the optimization process, resulting in higher management costs (or difficulty of changing these variables). From test data, we found that the total minimized cost to reduce churn probabilities of all customers by 20% is 1,329,440.73 (units of management costs, or difficulty of controlling efforts), ranging from 0 to 511.57 among customers.

	A	B	C	D	E	F	G	H
1	rev_Mean	totmrc_M	ovrrev_Me	vceovr_Me	rev_Range	totmrc_Ra	ovrrev_Ra	vceovr_Ra
2	30.455	29.998	5.28E-19	7.94E-24	0.92756	0.01464	-1.24E-23	6.20E-21
3	16.817	23.254	-5.74E-20	-6.68E-20	-2.52E-21	11.834	8.90E-21	8.58E-21
4	30.592	30.008	-5.48E-23	-1.18E-23	1.8275	0.01545	-8.17E-19	-2.66E-19
5	11.789	12.639	6.5359	6.5879	17.593	14.432	18.79	18.803
6	29.949	31.858	-1.79E-20	6.75E-20	6.70E-21	3.5093	1.41E-21	9.94E-20
7	40.361	42.834	9.00E-19	-1.32E-18	1.22E-20	0.64904	3.85E-21	6.40E-21
8	243.08	134.99	81.987	81.987	72.11	0.00167	38.85	38.85
9	57.547	69.955	0.09717	5.96E-21	49.768	0.17038	0.38913	-3.08E-21
10	50.236	50.036	2.53E-23	2.54E-22	0.97561	0.08724	2.59E-20	-4.71E-20
11	128.8	75.002	55.312	55.312	72.649	0.00376	64	64

T	U	V	W
Customer ID	ChurnProb(BO)	ChurnProb(AO)	Mgmt Cost
2000001	0.3344307	0.2675445	1.2118
2000002	0.6878101	0.5502481	112.27
2000003	0.3526896	0.2821517	1.2844
2000004	0.7526791	0.6021433	142.86
2000005	0.5472862	0.437829	42.616
2000007	0.4409892	0.3527914	13.286
2000009	0.03811693	0.03049355	0.25365
2000010	0.2363117	0.1890494	0.88153
2000011	0.3658637	0.292691	2.6906
2000012	0.08548163	0.0683853	0.3322

Figure 2. Values of optimized control variables, churn probabilities, and management costs

Remember that the cost unit used in this study does not directly bear the accounting or financial values, but it measures the difficulty of controlling variables. Therefore, it is still valuable if we can reduce the total management cost by adjusting the scope of marketing campaigns (i.e., targeting a subset of customers only) or changing the objective values of the primary marketing strategy based on PLS optimization (e.g., reducing churn probabilities of all customers by 10%, not 20%). While it is possible to reduce the costs further by changing objective values, it is a fairly straightforward process to implement another optimization problem by repeating the steps explained in previous Section 4.3. Therefore, we limit our interests to two supplementary approaches that can be combined with the PLS optimization process to further

reduce the cost by targeting a subset of customers. The first approach is to sort the customers based on their estimated probabilities of churning and target only customers with high churn probabilities to minimize the cost. Another supplementary approach is to target only customers with high low churn management costs, which will be discussed in the following Section. One of the nice properties of these supplementary approaches is that we do not need to solve another optimization problem because the objective function (i.e., reducing churn probabilities of all customers by 20%) is not changed at all and hence the same outputs of controllable variables before and after optimization can be used to compute the associated costs of the chosen subset of customers. We summarize objectives and expected benefits of primary and supplementary marketing strategies in Table 4.

Table 4. Objectives and expected benefits of marketing strategies

Churn Management Strategies	Marketing Objectives	Expected Benefits
Primary: PLS Optimization	Reduce churn probabilities of all customers by *x*% while minimizing management costs of all customers	Increase revenue by retaining more customers because of reduced churn probabilities
Supplementary: Customer Targeting	Target customers with high churn probabilities only	Reduce cost by targeting customer subsets
	Target customers with low churn management costs only	Reduce cost by targeting customer subsets

To target customers with high churn probabilities only, we first sort the customer records in the test data in the descending order in terms of their estimated churn probabilities after the PLS optimization model was applied. At this moment, each customer record is also associated with the estimated management cost to reduce its churn probability by 20%. Then marketing managers may select top *x*% of customers who are predicted to be most likely to churn and use two evaluation metrics, hit rate and cumulative percentage of management cost, to measure the success of marketing campaign. The hit rate is defined as the number of correctly identified churners out of churner candidates. When only *x*% of customers predicted most likely to churn are considered, it is called a hit rate at a target point *x*%. For example, if the model is required to select 1000 customers who are most likely to churn from 10,000 observations, and 100 of them turn out to be one of 500 actual churners, then a hit rate at target point 10% (1000/10,000=10%) is 20% (100/500=20%). At the same time, the cumulative percentage of management cost is also computed at a target point *x*% by dividing the sum of management costs of all customers up to a

target point *x*% by the total minimized management costs (1,329,440.73) of all the customers. The lift curve and the management cost curve show the trend of hit rates and the cumulative percentage of management costs over all possible target points, respectively. It is important to understand that a model with a higher lift curve above the diagonal line (i.e., a model with a higher accuracy than a random model) is preferred, while a model with a lower management cost curve under the diagonal line (i.e., a model with a lower cost than a random model) is preferred.

According to Figure 3, it is still possible to improve the outcome of marketing campaign by targeting only top *x*% of customers who are most likely to churn. The lift curve shows that the hit rates are higher across all target points than those of a random selection scenario (shown along the diagonal line). Note that this is an additional benefit of PLS control and optimization model to reduce churn probabilities of all customers to a certain level. We also notice that the management cost curve is located above the diagonal line. This indicates that when only top *x*% of customers who are most likely to churn are chosen and targeted, it will cost more (or be more difficult to manage) than the case of targeting randomly chosen customers. We attribute this finding to the fact that marketing managers need to control more controllable variables associated with higher management costs to reduce the churn probability for customers who are most likely to churn.

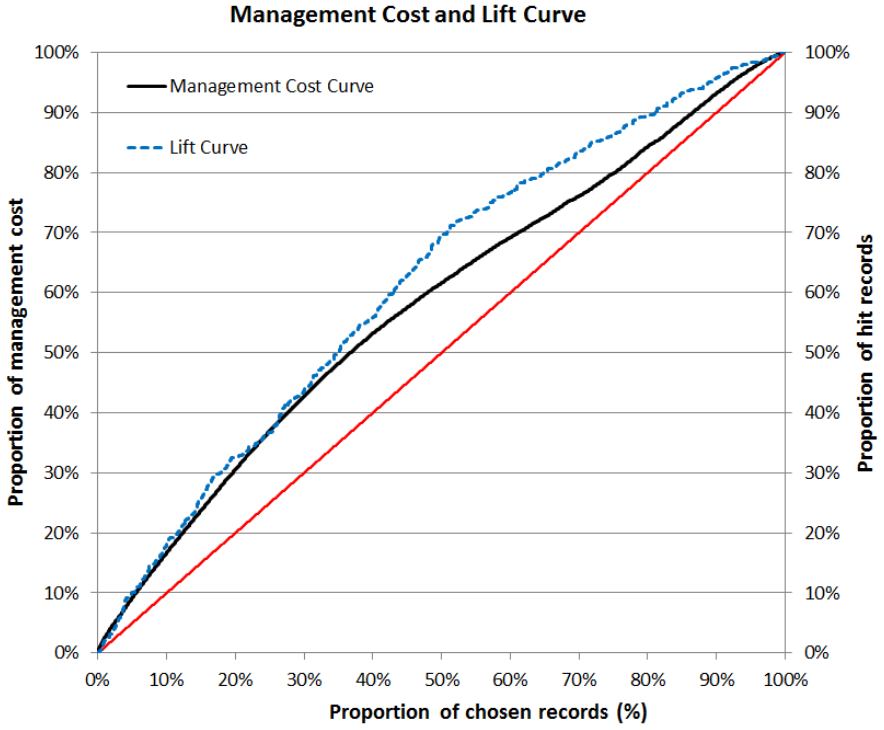

Figure 3. Management cost curve and lift curve based on churn probability

The final decision of selecting the best target point should consider both differences in management costs and hit rates of two models. For example, if the service provider believes that extra revenue generated by targeting only top *x*% of customers and predicting churners more accurately (i.e., measured by the difference between a lift curve and a diagonal line) is greater than extra management cost of targeting top *x*% of customers (i.e., measured by the difference between a management cost curve and a diagonal line), it is suggested that marketing managers target top *x*% of customers only based on their churn probability and enjoy supplemental financial benefit from customer targeting. Note, however, that even if marketing managers do not target a small portion of customers, they already accrue benefit of the PLS optimization model that sustain revenue of retained customer by reducing their churn probability by 20% during the optimization process.

5.2.2 Customer selection based management cost

Another customer selection method is to select candidate customers based on their management costs. In this scenario, we sort all the customers based on their management cost and select and target sequentially top *x*% of customers who are associated with the lowest management costs. We expect that the management cost of such a scenario will result in a cost curve under the diagonal line. In particular, the cost curve is assumed to take a shape of J in which the cumulative management cost of top *x*% of customers is very low, but it is increasing at an exponential rate once more customers with higher management costs are selected for churn management campaign. However, we expect that the lift curve is not significantly different from the diagonal line or even it is placed under the diagonal line because customers with lower management costs are not likely to be churners.

We present management cost curve and lift curve in Figure 4, and these curves confirm our expectation. First, the lift curve of the proposed supplemental marketing model is located under the diagonal line across the entire proportion of chosen customers except 80% or higher. For example, the proportion of hit records when top 50% of customers with the lowest management costs are targeted for churn management program is about only 40% of entire hit records (=619 records). This implies that randomly predicting churners is likely to identify churners slightly better than selecting customers in the ascending order of management costs and predicting them as possible churners. While this finding seems to be discouraging, it is well

expected because highly likely churners are associated with high management costs as observed in Section 5.2.1, but the proposed supplemental marketing model focuses on customers with low management costs only (i.e., less likely churners) to minimize management costs.

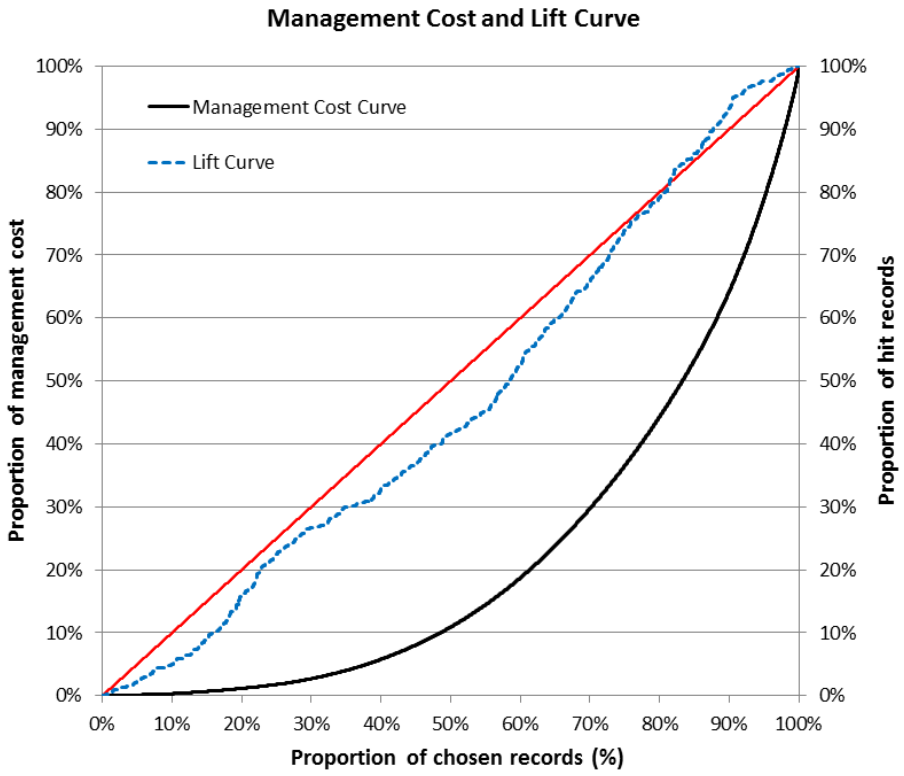

Figure 4. Management cost curve and lift curve based on management cost

In contrast, according to the J shape management cost curve, the cumulative management cost of top 50% of customers with the lowest management costs explain only 10% of the total management cost of all customers in the test data. In particular, the cumulative management costs of the proposed model over target points at 10%, 20%, and 30% of records are less than 3% of the total management cost of all customers, while those of a random model can occupy up to 30% of the total management cost. Note that the management cost curve of a random model is represented as a diagonal line. Our simple calculation shows that, up to target point 50%, the random model will cost between four times (at target point 50%) and ten times (at target point 10%) more than the proposed model that intelligently selects candidate customers based on associated management costs. Overall, we claim with confidence that the proposed supplemental model is still superior to a random model because it reduces management costs significantly compared with a random model while hit rates of two models are comparable. Note that the

proposed PLS model help marketing managers accomplish their marketing objectives not only by reducing the churn probability of all customers via optimization routine (i.e., primary churn management model) but also by reducing the management costs or hit rates via customer selection for target marketing (i.e., supplemental churn management model).

6. CONCLUSION AND FUTURE RESEARCH

In this paper, we present a churn optimization model based on the PLS prediction and control model. At first, a PLS regression model with chosen variables that meet a certain threshold is calibrated to estimate the churn probability of all the service users, and it is turned into an optimization model after dividing variables into controllable and uncontrollable variables, and associating management costs and other constraints with controllable variables. Then, the optimization problem is solved by using a SQP algorithm. The advantages of the proposed model are numerous. In terms of methodology, it significantly enhances the scalability and interpretability of a prediction model by including only predictive variables in its final model. In our example, it selects only 46 variables (out of 123 original variables) with high VIP scores and then constructs six latent variables of these 46 variables for final prediction. In addition, it allows marketing managers to explicitly develop a marketing campaign as a mathematical optimization model and efficiently solve it using a SQP algorithm. In particular, the proposed model and solutions are found to be robust to the dramatic changes in churner distributions of training and test data sets.

From the perspective of marketing managers, the proposed model is very customizable and generalizable by allowing managers to incorporate different marketing strategies and associated marketing variables in each marketing strategy. In particular, marketing managers can not only chose appropriate and preferred strategies (e.g., CMS, RMS, and DMS in this study) and controllable variables to meet their unique needs, but also apply their subjective weights and management costs to control variables. By incorporating a triangulation method to estimate relative management costs of each controllable variable, the proposed method not only allows different marketing managers to use different cost schemes but also the outcomes of the proposed model can be still applicable to different scenarios as long as the relative ranking of management costs for controllable variables is remained the same.

Finally, the managerial and financial outcomes of the proposed model can be discussed at two levels. First of all, the proposed model can reduce the churn probability of all the customers in the database while minimizing the cost of marketing campaign. The reduced churn probability of all the customers in turn implies that customers are less likely to terminate their current service contract and hence the service provider enhances the profitability over the life time of customers. Another advantage of the proposed model is that it allows marketing managers to accrue additional financial profits by limiting the scope of marketing campaigns to a subset of customers based on estimated churn probability or management cost.

One of limitations of the proposed model is that management costs associated with marketing variables do not directly bear the accounting or financial values and, hence, it is not straightforward to estimate the financial values of marketing campaigns. Therefore, a follow-up study may estimate the impact of input variables when they are associated with absolute financial values on the outcomes of the proposed model. In particular, it will be interesting to estimate the financial values from reduced churn probability of all the customers in the optimization process for a fixed structure of management costs of controllable marketing variables while changing the objective goal (e.g., how much should we reduce the churn probability of customer). At the same time, additional financial profits from customer selected based on churn probability or management cost can also be estimated. By doing so, marketing managers can not only determine an optimal objective goal in the optimization process but also maximize the financial profits for target marketing.

ACKNOWLEDGEMENTS

The author wishes to thank the Teradata Center for CRM at Duke University for making the data sets available. This research by Hyeseon Lee was supported with Basic Science Research Program through the National Research Foundation of Korea (NRF) from the Ministry of Education, Science and Technology (2010-0003628).

REFERENCES

[1] W.H. Au, K. Chan and X. Yao, A novel evolutionary data mining algorithm with applications to churn prediction. IEEE Transactions on Evolutionary Computation (7:6), 2003, pp. 532–545.

[2] W. Buckinx and D.V. Poel. (2005) Customer base analysis: partial defection of behaviorally loyal clients in a non-contractual FMCG retail setting. European Journal of Operational Research, 164: pp. 252–268.

[3] J.R. Bult and T. Wansbeek. (1995) Optimal selection for direct mail. Marketing Science, 14(4): pp. 378–394.

[4] P.Y. Chen and L.M. Hitt. (2002) Measuring switching costs and the determinants of customer retention in internet-enabled businesses: A study of the online brokerage industry. Information Systems Research, 13(3): pp. 255–274.

[5] I-G Chong, S.L. Albin, and C-H Jun. (2007) A data mining approach to process optimization without an explicit quality function. IIE Transactions, 39: pp.795–804.

[6] M. Cohn. Agile Estimating and Planning. Prentice Hall, 2006.

[7] K. Coussement and D. Van den Poel. (2008) Churn prediction in subscription services: An application of support vector machines while comparing two parameter-selection techniques. Expert Systems with Application, 34(1): pp. 313–327.

[8] A. Eshghi, D. Haughton, and H. Topi. (2007) Determinants of customer loyalty in the wireless telecommunications industry. Telecommunications Policy, 31(2): pp. 93–106.

[9] C. Fornell and B. Wernerfelt. (1987) Defensive marketing strategy by customer complaint management: A theoretical analysis. Journal of Marketing Research (24): pp. 337–346.

[10] P. Geldadi and B. Kowalski. (1986) Partial least-squares regression: A tutorial. Analytica Chemica Acta, 185: pp. 1–17.

[11] P. Gill, W. Murray, M. Saunders, and M. Wright. (1984) Procedures for optimization problems with a mixture of bounds and general linear constraints. ACM Transactions on Mathematical Software, 10(3): pp. 282–298.

[12] N. Glady, B. Baesens, and C. Croux. (2009). Modeling churn using customer lifetime value. European Journal of Operational Research, 197: pp. 402–411.

[13] F. Gönül and M. Z. Shi. (1998) Optimal mailing of catalogs: A new methodology using estimable structural dynamic programming models. Management Science (44(9): pp. 1249–1262.

[14] A. Gustafsson, M.D. Johnson and I. Roos. (2005) The effects of customer satisfaction, relationship commitment dimensions, and triggers on customer retention. Journal of Marketing, 69(4): pp. 210–218.

[15] J. Hadden, A. Tiwary, R. Roy, and D. Ruta. (2007) Computer assisted customer churn management: State-of-the-art and future trends. Computers & Operations Research, 34(10): pp. 2902–2917.

[16] R. Hagafors and B. Brehmer. (1983) Does having to justify one's decisions change the nature of the decision process? Oganizational Behavior and Human Performance, (31): pp. 223–232.

[17] P.M. Johnson, C.A. Moore, J.A. Dane, and R.S. Brewer. (2000) Empirically guided software effort estimation. IEEE Software, 17(6): pp. 51–56.

[18] M. Kim, M. Park and D. Jeong. (2004) The effects of customer satisfaction and switching barriers on customer loyalty in Korean mobile telecommunication services. Telecommunications Policy, 28(2): pp. 145–159.

[19] N. Kim, K.-H. Jung, J. Lee, and Y. Kim. (2012) Uniformly subsampled ensemble (USE) for churn management: Theory and implementation. *Expert Systems with Application*, 39(15): pp. 11839–11845.

[20] Y. Kim, W. N. Street, G. J. Russell, and F. Menczer. (2005) Customer targeting: A neural network approach guided by genetic algorithms. Management Science, 51(2): pp. 264–276.

[21] Y. Kim. (2006) Toward a successful CRM: Variable selection, sampling, and ensemble. Decision Support Systems, 41(2): pp. 542–553.

[22] V. Kumar and D. Shah. (2004) Building and sustaining profitable customer loyalty for the 21st century. Journal of Retailing, 80(4): pp. 317–329

[23] E.K. Laitinen. (2008) Data system for assessing probability of failure in SME reorganization. Industrial Management & Data Systems, 108(7): pp. 849–866.

[24] A. Lederer. (1998) A casual model for software cost estimating error. IEEE Transactions on Software Engineering, 24(2): pp. 137–148.

[25] A. Lederer and J. Prasad. (1992) Nine management guidelines for better cost estimating. Communications of the ACM, 35(2): pp. 51–59.

[26] H. Lee, Y. Kim, Y. Lee, H. Cho, and K. Im. (2011) Mining churning behaviors and developing retention strategies based on a partial least square (PLS) model. Decision Support Systems, 52(1): pp. 207–216.

[27] E. Malthouse, A. Tamhane, and R. Mah. (1997) Nonlinear partial least squares. Computers and Chemical Engineering, 21(8): pp. 875–890.

[28] E. Miranda. (2001) Improving subjective estimates using paired comparisons. IEEE Software, 18(1): pp. 87–91.

[29] V. Mittal and W. Kamakura. (2001) Satisfaction, repurchase intent, and repurchase behavior: Investigating the moderating effects of customer characteristics. Journal of Marketing Research, 38: pp. 131–142.

[30] M.C. Mozer, R. Wolniewicz, D.B. Grimes, E. Johnson and H. Kaushansky. (2000) Predicting subscriber dissatisfaction and improving retention in the wireless telecommunications industry. IEEE Transactions on Neural Networks, 11(3): pp. 690–696.

[31] S. A. Neslin, S. Gupta, W. Kamakura, J. Lu, and C. Mason. (2006). Defection detection: Improving predictive accuracy of customer churn models. Journal of Marketing Research, 43(2): pp. 204–211.

[32] W. Reinartz, J. Thomas, and V. Kumar. (2005) Balancing Acquisition and Retention Resources to Maximize Customer Profitability. Journal of Marketing, 69(1): pp. 63–79.

[33] R. Rosipal and L. J. Trejo. (2002) Kernel partial least squares regression in reproducing kernel Hilbert space. Journal of Machine Learning Research, 2: pp. 97–123.

[34] P.E. Rossi, R. McCulloch, and G. Allenby. (1996) The value of household information in target marketing. Marketing Science, 15(3): pp. 321–340.

[35] J. Shawe-Taylor and N. Cristianini. (2004) Kernel Methods for Pattern Analysis, Cambridge University Press.

[36] K. Schittowski. (1985) NLQPL: A FORTRAN-subroutine solving constrained nonlinear programming problems. Annals of Operations Research, 5: pp. 485–500.

[37] R. Venkatesan and V. Kumar. (2004) A customer lifetime value framework for customer selection and resource allocation strategy. Journal of Marketing, 68(4): pp. 106–125.

[38] S. Vicinanza, T. Mukhopadhyay, and M.J. Prietula. (1991) Software effort estimation: An exploratory study of expert performance. Information Systems Research, 2(4): pp. 243–262.

[39] M. C. Wiener, L. Obando and J. O'Neill. (2010) Building process understanding for vaccine manufacturing using data mining. Quality Engineering, 22(3): pp. 157–168.

[40] Y. Xiea, X. Lia, E.W.T. Ngaib and W. Yingc. (2009). Customer churn prediction using improved balanced random forests. Expert Systems with Applications, 36(3): pp. 5445–5449.